(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 0 535 242 B1

(12) **EUROPEAN PATENT SPECIFICATION**

(45) Date of publication and mention of the grant of the patent:
03.09.1997 Bulletin 1997/36

(21) Application number: **92907951.5**

(22) Date of filing: **18.03.1992**

(51) Int Cl.[6]: **C12Q 1/68**

(86) International application number:
PCT/RU92/00052

(87) International publication number:
WO 92/16655 (01.10.1992 Gazette 1992/25)

(54) METHOD AND DEVICE FOR DETERMINING NUCLEOTIDE SEQUENCE OF DNA

VERFAHREN UND VORRICHTUNG ZUR BESTIMMUNG VON DNS-NUKLEOTID-SEQUENZEN

PROCEDE ET DISPOSITIF DE DETERMINATION DE LA SEQUENCE DE NUCLEOTIDES D'ADN

(84) Designated Contracting States:
AT BE CH DE DK ES FR GB GR IT LI LU MC NL SE

(30) Priority: **18.03.1991 SU 4919321**

(43) Date of publication of application:
07.04.1993 Bulletin 1993/14

(73) Proprietor: **INSTITUT MOLEKULYARNOI BIOLOGII IM. V.A.ENGELGARDTA ROSSIISKOI AKADEMII NAUK**
Moscow, 117984 (RU)

(72) Inventors:
- **KHRAPKO, Konstantin Radievich**
  Moscow, 121433 (RU)
- **KHORLIN, Alexandr Anatolievich**
  Moscow, 117342 (RU)
- **IVANOV, Igor Borisovich**
  Dolgoprudny, 141700 (RU)
- **ERSHOV, Gennady Moiseevich**
  Moscow (RU)
- **LYSOV, Jury Petrovich**
  Moscow, 129344 (RU)
- **FLORENTIEV, Vladimir Leonidovich**
  Moscow, 103473 (RU)
- **MIRZABEKOV, Andrei Darievich**
  Moscow, 333775 (RU)

(74) Representative: **Ablewhite, Alan James**
MARKS & CLERK,
57/60 Lincoln's Inn Fields
London WC2A 3LS (GB)

(56) References cited:
EP-A- 0 132 621
EP-A- 0 159 719
EP-A- 0 266 787
EP-A- 0 322 311
WO-A-90/04652

- **DNA SEQUENCE, vol. 1, 1991, UK; K.R. KHRAPKO et al., pp. 375-388**
- **BIOLOGICAL ABSTRACTS, vol. 92, no. 7, 01 October 1991, Philadelphia, PA (US); P.A. PEVZNER et al., no. 71294**
- **FEBS LETTERS, vol. 256, no. 1/2, October 1989, AMSTERDAM (NL); K.R. KHRAPKO et al., pp. 118-122**

EP 0 535 242 B1

Printed by Jouve 75001 PARIS (FR)

Description

Field of the Invention

The present invention relates to molecular biology, and particularly to a method for determining DNA nucleotide sequence, and a device for carrying this out.

Background of the Invention

By now, several methods have been described for assaying a nucleotide sequence and identifying individual substitutions of bases by hybridization techniques (Cotton, R.G.H. // Biochem. J. 1989, V. 263, pp. 1-10).

The most widespread techniques are those in which a test DNA fragment is attached to a membrane and hybridized thereon with labelled oligonucleotides (Wallace, P.B., Shaffer, J., Murphy, R.F., Bonner, J., Hirose, T., Itakura, K. // Nucleic Acid Res., 1979, V. 6 pp. 3543-3557).

Known in the art are a method and a device for determining a DNA nucleotide sequence (E.Southern et al., WO-A- 8910977), which method comprises synthesizing oligonucleotides on a glass support, effecting hybridization with radioactively- or fluorescently-labelled test DNA, washing in the duplex dissociation conditions, detecting the presence of individual substitutions in the test sequence by analyzing the autoradiographic patterns or the intensity of fluorescence at individual dots, and reconstructing the DNA nucleotide sequence on the basis of data analysis. A device for carrying out said method comprises a supporting film or glass plate and a matrix covalently attached to the surface thereof, the matrix comprising the whole set or a selected part of oligonucleotides of desired length, the latter oligonucleotides being capable of taking part in the hybridization reactions. The surface of the support to which the oligonucleotides are attached is made of glass. Another method of DNA sequencing is based on hybridisation of a DNA fragment with a complete set of fixed length oligonucleotides which are immobilised individually as dots on a matrix. The sequence of the DNA fragment being determined on the basis of the dissociation curves of the immobilised duplexes (K. R. Khrapko et al, FEBS Lett, 256, 118-122, 1989).

The above methods and devices, however, have low sensitivity.

The value of the signal from the labelled DNA (laid out in the near-surface layer of the support) obtained from each individual matrix element is limited by the substrate surface capacity (as regards the covalently attached oligonucleotides) and it cannot be raised without increasing either the area of the matrix element or the sensitivity of the labelling marker. These limitations reduce the resolving power of the method and device, make it difficult to miniaturize the matrix, increase the requirements imposed on the sensitivity and resolving power of the detector, and raise the consumption of the reagents. The method is rather complicated because, even in assaying a test DNA fragment, it requires a series of successive hybridizations to be performed, with additional rounds of oligonucleotide matrix synthesis to be performed at one-letter step, for each dot where hybridization has not yielded unambiguous information on the sequence and, therefore each time new optimal hybridization conditions (temperature, reagent concentrations, etc.) have to be chosen, which involves further experimentation and considerable expenditure of time and reagents.

Disclosure of the Invention

It is an objective of the invention to alter the method and apparatus in such a way that to improve their efficiency, to increase sensitivity, accuracy and reproducibility, to simplify recognition of point mutations in the nucleotide sequence, and to reduce the expenditure of reagents.

In the present method for determining the DNA nucleotide sequence which comprises formation of an array of oligonucleotides, their hybridization with the labelled test DNA, washing under duplex dissociation conditions, identification of every single substitution of bases in the test DNA by analyzing distribution of the labelling marker, and eventual computerized reconstruction of the DNA nucleotide sequence, the above objective is achieved by that the invented method involves formation of an array of oligonucleotides at such concentrations thereof which ensure the same temperature of duplex dissociation for all fully complementary duplexes.

To achieve reliable discrimination of fully complementary duplexes from duplexes having point mutations (mismatches), it is preferred to carry out the washing step at a fixed temperature gradient.

In order to improve accuracy and reduce the duration of data analysis it is preferred during washing to record the dependence of the amount of the remaining duplexes on the temperature to be compared with the dependence for the known sequence of perfect double helical DNA.

To reduce the duration of the analysis, it is advantageous to form an array of oligonucleotides at concentrations allowing the hybridization and washing of fully complementary duplexes to be carried out at the same temperature.

The present method makes it possible to considerably simplify the procedure as compared to the prior art method, and to improve its sensitivity, accuracy and reproducibility. The present method helps to improve the efficiency of the

procedure by making it more economic in terms of time, labour, and reagents used. Furthermore, the size of the cells and the interstices between them are chosen in such a way as to conveniently combine the apparatus of the invention with the existing technological and measuring equipment.

The objective of the invention is further achieved by that the device for determining the DNA nucleotide sequence comprises a solid support and a matrix including an array of oligonucleotides of desired length, the matrix being attached, according to the invention, to the support by means of a gel layer having a thickness of no more than 30 μm. Preferably, the gel layer consists of a set of spaced "dots", according to the number of matrix elements. The gel layer provides for a three-dimensional attachment of the oligonucleotides at a capacity exceeding considerably the capacity of a molecular layer, while the gel layer comprising a plurality of spaced dots allows the desired number of oligonucleotides to be localized within a selected gel volume. All this makes it possible to miniaturize the matrix, and raise the rate of all the processes, and thus reduce the duration of the procedure, and improve the sensitivity, resolving power, accuracy and reproducibility of the method and device, and lower the consumption of the reagents.

The present device is a device, in which the surface of each portion of the gel layer ("cell") has the shape o a square with a side length of 25 to 100 μm and the spacing between the squares equal to double the side length. This design permits a dynamic aquilibrium to be rapidly established during the hybridization process, the consumption of reagents to be lowered, the sensitivity raised, and non-toxic, non-radioactive markers to be used.

In all the embodiments of the device, it is preferred to make the layer of a polyacrylamide gel which is convenient in use, readily available and reproducible.

The present device helps to simplify the procedure for determining a DNA nucleotide sequence, to reduce its duration, to improve the sensitivity,accuracy and reproducibility, and to lower the consumption of the reagents.

A Brief Description of the Drawings

The invention is explained below by a detailed description of its embodiments with reference to the accompanying drawings, as follows.

Fig. 1 is a scheme of the apparatus for determining the DNA nucleotide sequence as viewed from above.

Fig. 2 is a longitudinal sectional view of F 1;

Fig. 3 is a scheme of chemical reactions occurring during immobilization of an oligonucleotide in a polyacrylamide gel.

Fig. 4 shows the washing curves of AT-rich duplexes (a) and GC-rich duplexes (b), with the amount of remaining duplexes plotted on the Y-axis, in %, and the washing temperature on the X-axis, in °C Fig. 5 presents the dependence of the duplex washing temperature (the duplex shown at top, M stands for matrix) on the concentration of immobilized oligonucleotide. The remaining duplexes are plotted on the Y-axis, in %, and the washing temperature on the X-axis, in °C ;

Fig. 6 shows the dependence of the washing temperatures of AT-rich and GC-rich duplexes on the concentration of oligonucleotides immobilized in gel. The amount of remaining duplexes, in %, are plotted on the Y-axis, and the washing temperature, in °C, on the X-axis;

Fig. 7 is a comparative diagram for discriminating mismatches in duplexes of different GC composition on a matrix having selected concentrations of immobilized oligonucleotides.

Fig. 8. The scheme of micro-matrix illustrates the specificity of hybridization and sensitivity of the method.

Best Mode of Carrying Out the Invention

The present device for determining DNA nucleotide sequence comprises support 1 (Figs. 1 and 2), preferably a glass plate, and matrix 2 attached to the surface thereof by means of a gel layer of less than 30 μm thick. The gel layer may comprise a multitude of portions 3, according to the number of elements in matrix 2 , spaced from one another by interstices 4. The interstices 4 may have different dimensions. The gel portions 3 may have different shapes.

A preferred device is one in which each portion 3 has the shape of a square with a side length of 25 to 100 μm and interstices 4 between the squares are equal to double the side length. The layer can be made of various gels, preferably a polyacrylamide gel.

The matrix can be manufactured as follows;

Two slides, one of which is pretreated with Bind Silane, and the other with Repel Silane, the latter slide was lubricated with a thin layer of Triton® X-100. The slides are stacked with the use of spacers of less than 30 μm thick, the resulting spacing between them is filled with a gel solution and the gelling process is allowed to complete, whereupon the top slide is removed. The gel-coated lower slide is dried, part of the gel is removed, for example mechanically, so that the gel portions (cells) separated by interstices remain on the slide surface. The surface thus obtained is treated for 2 to 5 minutes with Repel Silane, washed first with alcohol and then with bidistilled water, and dried. Oligonucleotides

containing 3-methyluridine at the 3'-end are oxidized with 1 mM of sodium periodate for 10 minutes to 1 hour at room temperature, precipitated with 10 volumes of 2% $LiClO_4$ in acetone, and dissolved in water. Then, the oligonucleotides are immobilized in the gel. For this purpose, the cells of the air-dried matrix are filled with microdoses of oxidized oligonucleotides of identical volume (one type into one cell) from the available stock.

The array of the oligonucleotides is formed in such a way that their concentrations ensure the same dissociation temperature for all fully complementary duplexes or at concentrations allowing subsequent hybridizations and washing of fully complementary duplexes to be carried out at the same temperature. A test DNA fragment labelled with radioactive or fluorescent marker in a buffer solution is applied to the matrix with the preformed array of oligonucleotides (in the process, the solution fully covers all areas containing immobilized oligonucleotides). The array of oligonucleotides is then hybridized with the added labelled test DNA, and the duplexes are washed away in dissociation conditions. Single substitutions of bases in the test DNA are identified by analyzing distribution of the marker.

The test DNA sequence is reconstructed on the basis of data analysis. To reliably distinguish fully complementary duplexes with point mutations (mismatches), duplex washing in the present method is effected at a fixed temperature gradient. In order to reduce duration and improve accuracy of this analysis, it is preferred that in the course of washing the dependence of the amount of remaining duplexes on temperature should be determined and compared with the dependence for the known DNA sequence of perfect duplexes. It has been established that nearly always a temperature could be found at which the ratio of hybridization signals from a fully complementary duplex and a corresponding duplex containing point mutations is sufficiently high (at least 10:1) to reliably distinguish one from the other. The exception is provided by some terminal mismatches which may have a high stability. This does not, however, restrict the scope of the present method. Indeed, dealing with known sequences (for example, detecting mutations in them) it is always possible to select for immobilization such an oligonucleotide in which the expected base substitution would be located within the duplex. On the other hand, in analysis of an unknown nucleotide sequence (for example, in DNA sequencing), the problem of terminal mismatches can be easily solved by sacrifying some information while processing the data on hybridization between a DNA fragment and an oligonucleotide matrix by means of a computer. Our calculation methods are characterized by a high stability owing to excess information.

For a better understanding of the present invention, some aspects of its actual realization will be exemplified in the following.

Example 1

Manufacture of a matrix with an array of oligonucleotides:

Oligonucleotides are synthesized by a solid-phase phosphoramidite method (protection is removed in a saturated aqueous ammonia solution at 55°C for 12 hours) and purified by electrophoresis in a polyacrylamid gel. Oligonucleotides were labelled by ([γ-^{32}P]ATP by polynucleotide-kinase T4) at the 5'-end to achieve a specific activity of 3 μCi/pmole.

Two slides, one of which pretreated with Bind Silane and the other with Repel Silane (LKB) and lubricated with a thin layer of Triton X-100 are positioned at a distance from one another by means of 30 μm thick spacers. The resulting clearance between the slides is filled with 8% acrylamide solution, 30:1 N,N'-methylene bis-acrylamide, ammonium persulfate and TEMED, which is allowed to polymerize for 1 hour. As a result, a 30 μm thick gel layer is formed between the slides, the dimensions of the layer being determined by the size of the slides. When the polymerization is complete, the upper slide is removed. The lower slide coated with a polyacrylamide layeris treated with 50% hydrazine for 1 hour at room temperature.

Oligodesoxynucleotides containing 3-methylyl uridine at the 3'-end are oxidized with 1 mM sodium periodate for 1 hour at room temperature, precipitated with 10 volumes of 2% $LiClO_4$ in acetone and dissolved in water. The air-dried matrix ready for immobilization is covered dropwise by means of a micromanipulator quipped with a capillary tip dispenser with 0.5 μl droplets of oxidized oligonucleotide (at a concentration of 10 pmol/μl). The plates are then exposed for 4 hours in a moist chamber, dried for 0.5 hour in open air, washed with a hybridization buffer (1M NaCl, 10 mM Na phosphate, pH 7.0, 1 mM ethylene diamine tetra-acetic acid), flushed with water and stored dry at -20°C. The oligonucleotides are immobilized in cells of a square matrix.

The linker is represented by 3-methyluridine attached by a 5'-3' internucleotide phospho-diester bond to the oligonucleotide to be immobilized. 3-methyluridine was chosen due to the fact that it does not form strong hydrogen bonds with any natural bases. The scheme of the chemical reactions occurring in polyacrylamide gel during oligonucleotide immobilization is shown in Fig. 3.

The oxidation of the 3'-terminal ribonucleoside of oligonucleotide 1 (Fig. 3). with $NaIO_4$ produces a derivative 2 carrying a dialdehyde group at the 3'-end. On the other hand, when polyacrylamide 3 is treated with hydrazine, part of the amide groups are replaced with hydrazide groups 4, which readily react with 3'-dialdehyde, producing a relatively stable morpholine derivative 5.

The course of immobilization is monitored by the (5'-^{32}P] marker introduced by means of a kinase into the oligo-

nucleotides to be immobilized. The yield of immobilization (that is, the proportion of oligonucleotides irreversibly bound to the gel) is 80%. At the same time. the yield for the unoxidized oligonucleotide used as a nonspecific sorption control does not exceed 2%. Therefore, the proportion of molecules bonded specifically through their 3'-end is 98%.

The bond between an oligonucleotide and polyacrylamide is stable enough for the matrix to withstand at least 5 to 7 hybridization/washing cycles without any noticeable change in its hybridizing properties. The half-life of the oligonucleotide-gel bond at 60°C is 2 hours, and at 25°C, 36 hours.

The capacity of a carrier is evaluated by immobilizing the same amount of ^{32}P-labeled oligonucleotide diluted by unlabeled oligonucleotide to various specific activities. 100 pmoles of a cold oligonucleotide per one dot (i.e., per 1 mm^2 of surface area or 0.03 mm^3 of gel volume) do not saturate the bonding. Similar experiments with an oxidized periodate [α-^{32}P] UTP have shown that the gel capacity is equal to about 1 nmole per 1 mm^2 of the gel surface, which corresponds to the 30 mM concentration of active groups (concentration of amide groups is 1 M in 8% polyacrylamide).

Example 2

Reconstruction of the nucleotide sequence of a 17-membered desoxyoligonucleotide

Hybridization of four heptadekanucleotides of the phage M13 sequence primer: 5'-d(GTAAAACGACGCCAGT) and its three derivatives differing by one base (underlined):

5'-d(GTAAAACGATGGCCAGT),

5'-d(GTAAAACGAAGGCCAGT,

and

5'-d(GTAAAACGACGGCCAGT),

with immobilized oligonucleotides (comprising 7, 8, 9, 12 or 15 monomer units) fully or partially complementary to different portions of heptadekamers is effected on an oligonucleotide matrix prepared as described in Example 1.

The labelled DNA fragment (0.01. μCi, 30 fmoles) in 1 μl of a hybridization buffer (1 M NaCl, 10 mM of Na phosphate, pH 7.0, 1 mM ethylenediamine tetraacetic acid) is applied to a matrix of immobilized oligonucleotides so that each droplet of the hybridized mixture accurately covers the spot of the immobilized oligonucleotide, and incubated for 1 hour at 0°C. The matrix is flushed with the hybridization buffer at 0°C and then washed 10 times, 1 minute each, with 20 ml of the same buffer at a temperature risen by 5°C at each washing step. Upon each step, the hybridization signal is registered in each cell of the matrixe through a lead collimator with a radioactivity counter (Minimonitor 125, Victoreen, USA) equipped with a pulse adder.

The ratio of residual radioactivity to the starting radioactivity at a given point is plotted on a logarithmic scale versus temperature (Fig. 4).

Fig. 4 shows the washing curves of duplexes formed by the M13 primer or its analogous with GC- (Fig. 4a) and AT-rich (Fig. 4b) octanucleotides immobilized in gel, complementary to two different portions of the M13 primer, but producing defective duplexes with its derivatives (all the duplexes are shown in Fig. 4).

As the indicator of the duplexe stability, the washing temperature (T_w) is chosen so that the hybridization signal at the corresponding point decreases by a factor of 10 compared with the starting level. The washing curves of the duplexes were determined for different volume concentrations of the immobilized oligonucleotide in gel. As is clear from Fig. 5, T_w of a duplex depends strongly on the quantity of oligonucleotide immobilized in the spot of a given size. In the studied concentration range, the duplex washing temperature rises, within the experiment error, by a certain number of degrees when the immobilized oligonucleotide concentration becomes several times higher. This rule holds true for all the duplexes studied. In this way. the stability of each oligonucleotide can be varied. In this example, the concentration of 5 pmole per point keeps the duplexes stable within the range of 20 to 40°C.

In addition to the oligonucleotides shown in Fig. 4, a large number of 7-, 8-, 9-. 12- and 15-membered substances, fully or partially complementary to the M13 primer, and duplexes containing other mismatch types, have been studied. The results of washing the duplexes of heptadekadesoxynucleotides with immobilized octadesoxynucleotides are shown in the Table.

As is clear from Fig. 4 and the Table there is almost always a temperature at which the ratio of hybridization signals of a fully complementary duplex and a corresponding defective duplex is sufficiently high (at least 10) to reliably distinguish them. The exception is some terminal mismatches with abnormally high stability.

In the case of hybridization with octanucleotides, a sevenfold excess of information is obtained. Any nucleotide of a DNA fragment forms an inner pair with an immobilized oligonucleotide in six out of eight instances, whereas only in two instances it produces an end pair. A comparison is made between the washing curves to distinguish fully complementary duplexes from duplexes carrying point mutations and in this way single substitutions of bases in the DNA are detected. An initial heptadekamer is reconstructed from the overlapping of fully complementary duplexes.

Example 3

The dependence of the duplex washing temperature on the concentration of an immobilized oligonucleotide can be determined as follows.

The procedure is similar to that of Example 1. The dependence of the duplex washing temperature T_w on the concentration of an immobilized oligonucleotide is determined in a spot of 0.03 mm^3 volume within the oligonucleotide concentration range of 5.0, 1.5, 0.5 and 0.15 pmole.

The dependence of duplex washing curves on the volume concentration of an immobilized oligonucleotide in gel is shown in Fig. 5 (spot volume 0.03 mm^3, oligonucleotide concentration 5.0 (I), 1.5 (II), 0.5 (III) and 0.15 (IV) pmole). As is clear from the graph in Fig. 5, T_w of a duplex depends heavily on the amount of oligonucleotides immobilized in a spot of given size. For the concentration range studied, the duplex washing temperature increases, within the experimental error, by a definite number of degrees while the concentration of immobilized oligonucleotide rises by a definite number of times. This rule holds for all the duplexes studied (Fig. 5 shows one example).

Table

Thermal Dissociation of Perfect Duplexes and Duplexes Containing Single Mismatches

Perfect Duplexes (T_w, °C)	Mismatch GT (ΔT^*_w, °C)	Mismatch GA (ΔT^*_w, °C)	Mismatch CC (ΔT^*_w, °C)
GTCGTTTT[M] 3'– –GCAGCAAAAT– (28±0,5)	G_TTCGTTTT[M] 3'– –G$_T$AGCAAAAT– (-7±1)	G_ATCGTTTT[M] 3'– –G$_A$AGCAAAAT– (-16±2)	
CGTCGTTT[M] 3'– –GGCAGCAAAA– (28±0,5)	C^{G}TCGTTT[M] 3'– –GG$_T$AGCAAAA– (-13±2)	C^{G}TCGTTT[M] 3'– –GG$_A$AGCAAAA– (-21± 1)	C_CGTCGTTT[M] 3'– –G$_C$CAGCAAAA– (-3±0,5)
CCGTCGTT[M] 3'– –CGGCAGCAAA– (33±0,5)	CC^{G}TCGTT[M] 3'– –CGG$_T$AGCAAA– (-12±1)	CC^{G}TCGTT[M] 3'– –CGG$_A$AGCAAA– (-24±1)	C^{C}GTCGTT[M] 3'– –CG$_C$CAGCAAA– (-30±1)
GCCGTCGT[M] 3'– –CCGGCAGCAA– (41±0,5)	GCC^{G}TCGT[M] 3'– –CCGG$_T$AGCAA– (-11±1,5)	GCC^{G}TCGT[M] 3'– –CCGG$_A$AGCAA– (-25±1)	GC^{C}GTCGT[M] 3'– –CCG$_C$CAGCAA– (-36±2)
GGCCGTCG[M] 3'– –ACCGGCAGCA– (50±0,5)	GGCC^{G}TCG[M] 3'– –ACCGG$_T$AGCA– (-14±1,5)	GGCC^{G}TCG[M] 3'– –ACCGG$_A$AGCA– (-23±1)	GGC^{C}GTCG[M] 3'– –ACCG$_C$CAGCA– (-33±1)
TGGCCGTC[M] 3'– –GACCGGCAGC– (41±0,5)	TGGCC^{G}TC[M] 3'– –GACCGG$_T$AGC– (-12±0,5)	TGGCC^{G}TC[M] 3'– –GACCGG$_A$AGC– (-17±0,5)	TGGC^{C}GTC[M] 3'– –GACCG$_C$CAGC– (-37±3)
CTGGCCGT[M] 3'– –TGACCGGCAG– (36±0,5)	CTGGCC^{G}T[M] 3'– –TGACCGG$_T$AG– (-10±1,5)	CTGGCC^{G}T[M] 3'– –TGACCGG$_A$AG– (-6±1)	CTGGC^{C}GT[M] 3'– –TGACCG$_C$CAG– (-35±0,5)
ACTGGCCG[M] 3'–TGACCGGCA– (39±1)	ACTGGCC^G[M] 3'–TGACCGG$_T$A– (-5±1)	ACTGGCC^G[M] 3'–TGACCGG$_A$A– (-2±0,5)	ACTGGC^{C}G[M] 3'–TGACCG$_C$CA– (-30±0,5)

* Decrease in the washing temperature of a defective duplex compared with a perfect one

Note: Mismatches are given in bold type, *M* stands for matrix, and symbol "d" is ommited for simplicity

Example 4

The stability of fully complementary duplexes can be equalized as follows:

The procedure is the same as in Example 2, but the concentration varies for different immobilized oligonucleotides: 90 pmoles for AT, and 0.3 pmole for GC.

The dependence of stability (dissociation temperature) of duplexes on the concentration of immobilized oligonu-

cleotides in gel makes it possible to detect mismatches in duplexes of different GC composition "on a single plate". In Fig. 4, the two shown immobilized oligonucleotides are very different in their GC content, and as a result the fully complementary duplex of the AT-rich oligonucleotide (Fig 4a, curve I) is essentially less stable than the duplex of the GC-rich oligonucleotide containing mismatches (Fig. 4b, curves II and III). It must be clear that in this situation hybridization at a fixed temperature of a DNA fragment with a matrix containing both oligonucleotides will not allow to tell, within an accuracy of one base, whether or not this fragment has sequences complementary thereto.

The dependence of T_w on the immobilized oligonucleotude concentration can be used to equalize dissociation temperatures of duplexes of different GC composition. As shown in Fig. 6, the washing curves of the AT- and GC-rich duplexes (ΔT_w=30°C at the same concentration) can be brought into coincidence by properly selecting the concentrations of immobilized oligonucleotides at respective points.

One can equalize the washing curves of any array of oligonucleotides and thus obtain a "normalized" oligonucleotide matrix. The washing temperatures of fully complementary (perfect) duplexes for all the points of such a matrix are close to one another, therefore a single washing at an optimal temperature will suffice to unambiguously determine the points of the matrix where the DNA fragment under analysis has formed perfect duplexes.

Example 5

The hybridization and washing steps can be performed at the same temperature on a "normalized" matrix containing preselected concentrations of immobilized oligonucleotides, as follows:

The procedure is carried out similarly to Example 4.

The "normalized" matrix of two oligonucleotides (three points for each nucleotide) is hybridized and washed at 35°C. The principle can be understood from the comparative diagram in Fig. 7. Comparison of the residual signals indicates unambiguously the points where the duplexes were fully complementary, even though one terminal mismatch, GT, is very stable.

Example 6

Sensitivity, accuracy and reproducibility of the present method and device are illustrated as follows:

An oligonucleotide matrix is prepared as described in Example 1, except that prior to treatment with 50% hydrazine the glass plate coated with a gel layer is air-dried and part of the gel is removed, for example mechanically, to form squares having the side of 25 to 100 μm spaced from one another by interstices of 50 to 200 μm, respectively.

Hybridization is carried out with fragments labelled at the 3' end by fluorescent tetramethylrhodamine marker, which is introduced with the help of a terminal polynucleotide transferase (comparison of the respective washing curves showed tetramethylrhodamine to exert no influence on the duplex stability).

The use of a fluorescent marker allows hybridization to be carried out on a micro-scale. Fig. 8 shows a scheme of a micro-matrix, in which an octanucleotide 5'd(GGCCGTGG) is immobilized in gel squares having sides of 100 μm with 200 μm interstices between them. As can be seen in Fig. 8, a, hybridization with this matrix makes it possible to reliably detect substitutions of individual bases in the sequence (in Fig. 8), in square 1 (perfect duplex), and in squares 2, 3 and 4 (duplexes with mismatches GA, GT, and CC, respectively).

Similar experiments were carried out on micro-matrices, in which gel squares had the side of 25 μm, 30 μm, 50 μm, and 75 μm, with the interstices between them of 50 μm, 60 μm, 100 μm, and 140 μm, respectively.

Since distribution of the fluorescent marker can be measured at a very high sensitivity and spatial resolution (or example, with the aid of fluorescent microscope), detection of hybridization signals is limited only by the signal/background ratio and does not depend on the size of the object, whose fluorescence intensity is measured. Therefore, the sensitivity is inversely proportional to the area of the object(= oligonucleotide matrix cell). Miniaturization of the matrix allows the sensitivity of the method to be improved. For example, the squares "d" in Fig. 8, contain, respectively, 1 fmole, 100 amole and 10 amole of the tetramethylrhodamine derivative of desoxyuridine. The signal/background ratio for the square "d3" is equal to 2, which is sufficient to reliably assess the quantity of the substance.

Industrial Applicability

The claimed method and device can be used in medicine, molecular biology, and agriculture for the purposes of genetic diagnostics, DNA sequencing and mapping. and mutation detection

Claims

1. A method for determining DNA nucleotide sequence comprising the following steps:

- formation of an array of oligonucleotides wherein the oligonucleotides concentrations are chosen such as to ensure the same dissociation temperature for all fully complementary duplexes;
- hybridization of said oligonucleotides array with a test DNA labelled with a marker ;
- washing in conditions ensuring duplex dissociation;
- discrimination of single-base substitutions in the test DNA by analysing the distribution of the marker;
- reconstruction of the test DNA nucleotide sequence on a basis of data analysis.

2. A method as disclosed in claim 1, charactarized in that an array of oligonucleotides is formed at such concentratons which ensure that hybridization and washing of the fully complementary duplexes can be carried out at the same temperature in the course of washing.

3. A method as disclosed in claim 1 or 2, characterized in that washing is effected at a fixed temperature gradient.

4. A method as disclosed in any of claims 1 to 3, wherein the dependence of the amount of remaining duplexes on temperature is determined in the course of washing and compared with the corresponding dependence for a DNA of a known sequence.

5. A device for determining DNA nucleotide sequence, comprising a solid support (1) and a matrix (2) bound to said support and bearing an array of oligonucleotides of given length, characterized in that the matrix (2) is attached to the support (1) by a gel layer of a thickness not exceeding 30 µm and comprised of a multiplicity of gel portions (3), according to the number of matrix elements, which are separated from one another by interstices (4), each gel portion being a square of a 25-100 µm in length, with the interstices between the squares being twice the length of each square.

6. A device as claimed in claim 5, characterized in that the gel layer is made of a polyacrylamide gel.

Patentansprüche

1. Verfahren zum Bestimmen der DNA Nukleotidsequenz, umfassend die folgenden Stufen:

   Bildung einer Reihe von Oligonukleotiden, wobei die Oligonukleotidkonzentrationen so gewählt sind, daß die gleiche Dissoziationstemperatur für alle vollständig komplementären Doppelstränge gewährleistet ist,

   Hybridisierung der Oligonukleotidreihe mit einer mit einem Marker markierten Test DNA,

   Waschen bei Bedingungen, die Doppelstrangdissoziation gewährleisten,

   Unterscheidung von Einzelbasen-Substitutionen in der Test DNA durch Analysieren der Verteilung des Markers,

   Rekonstruktion der Test DNA Nukleotidsequenz auf einer Basis von Datenanalyse.

2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß eine Reihe von Oligonukleotiden bei derartigen Konzentrationen gebildet wird, die gewährleisten, daß Hybridisierung und Waschen der vollständig komplementären Doppel stränge bei der gleichen Temperatur im Verlauf des Waschens durchgeführt werden kann.

3. Verfahren nach Anspruch 1 oder 2, dadurch gekennzeichnet, daß Waschen bei einem festgelegten Temperaturgradienten bewirkt wird.

4. Verfahren nach einem der Ansprüche 1 bis 3, wobei die Abhängigkeit von der Menge verbleibender Doppel stränge von der Temperatur im Verlauf des Waschens festgelegt und mit der entsprechenden Abhängigkeit für eine DNA mit einer bekannten Sequenz verglichen wird.

5. Eine Vorrichtung zum Bestimmen von DNA Nukleotidsequenz, umfassend einen festen Träger (1) und eine Matrix (2), die an den Träger gebunden ist und eine Reihe von Oligonukleotiden mit gegebener Länge trägt, dadurch gekennzeichnet, daß die Matrix (2) an den Träger (1) durch eine Gelschicht mit einer Dicke, die 30 µm nicht übersteigt, angeheftet ist und aus einer Vielzahl von Gelteilen (3) gemäß der Anzahl von Matrixelementen zusam-

mengesetzt ist, welche voneinander durch Zwischenräume (4) getrennt sind, wobei jedes Gelteil ein Quadrat von 25-100 µm in Länge ist, wobei die Zwischenräume zwischen den Quadraten zweimal die Länge jedes Quadrats sind.

6. Eine Vorrichtung nach Anspruch 5, dadurch gekennzeichnet, daß die Gelschicht aus einem Polyacrylamidgel hergestellt ist.

Revendications

1. Procédé de détermination d'une séquence de nucléotides d'ADN comprenant les étapes suivantes:

   formation d'une batterie d'oligonucléotides dans laquelle les concentrations des oligonucléotides sont choisies de sorte qu'elles assurent la même température de dissociation pour toutes les doubles hélices entièrement complémentaires;

   hybridation de ladite batterie d'oligonucléotides avec un ADN à tester marqué par un marqueur;

   lavage dans des conditions assurant la dissociation des doubles hélices;

   discrimination des substitutions simples de bases dans l'ADN à tester par analyse de la répartition du marqueur;

   reconstruction de la séquence nucléotidique de l'ADN à tester sur la base de l'analyse des données.

2. Procédé selon la revendication 1, caractérisé en ce qu'une batterie d'oligonucléotides est formée à des concentrations telles qu'elles garantissent la possibilité d'effectuer l'hybridation et le lavage des doubles hélices entièrement complémentaires à la même température pendant le lavage.

3. Procédé selon la revendication 1 ou 2, caractérisé en ce que le lavage est effectué à un gradient de température fixé.

4. Procédé selon l'une quelconque des revendications 1 à 3, dans lequel la relation de dépendance entre la quantité de doubles hélices restantes et la température est déterminée au cours du lavage et comparée à la relation de dépendance correspondante pour un ADN de séquence connue.

5. Dispositif de détermination d'une séquence de nucléotides d'ADN, comprenant un support solide (1) et une matrice (2) liée audit support et porteuse d'une batterie d'oligonucléotides ayant une longueur donnée, caractérisé en ce que la matrice (2) est fixée au support (1) par une couche de gel d'une épaisseur ne dépassant pas 30 µm et est constituée d'une multitude de portions de gel (3), selon le nombre d'éléments de la matrice, qui sont séparées les unes des autres par des interstices (4), chaque portion de gel constituant un carré de 25 à 100 µm de longueur, les interstices entre les carrés ayant une longueur égale à deux fois celle de chaque carré.

6. Dispositif selon la revendication 5, caractérisé en ce que la couche de gel est faite d'un gel de polyacrylamide.

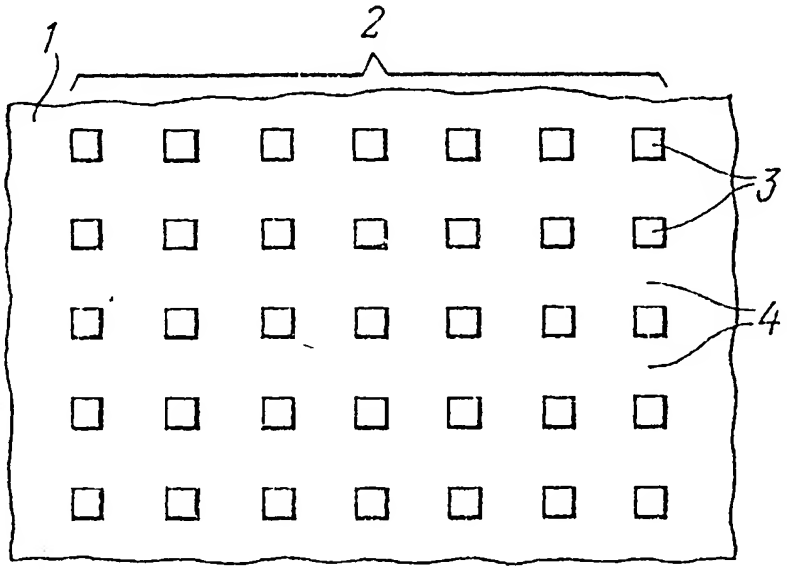


FIG.1

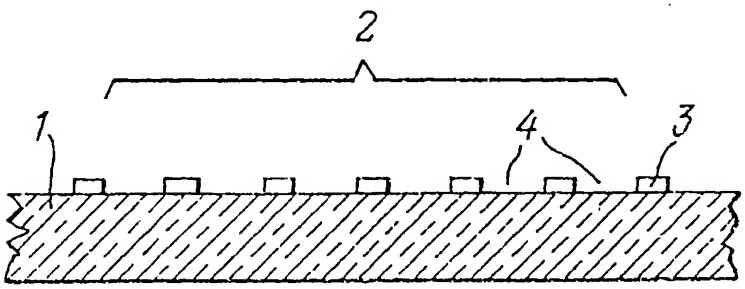


FIG.2

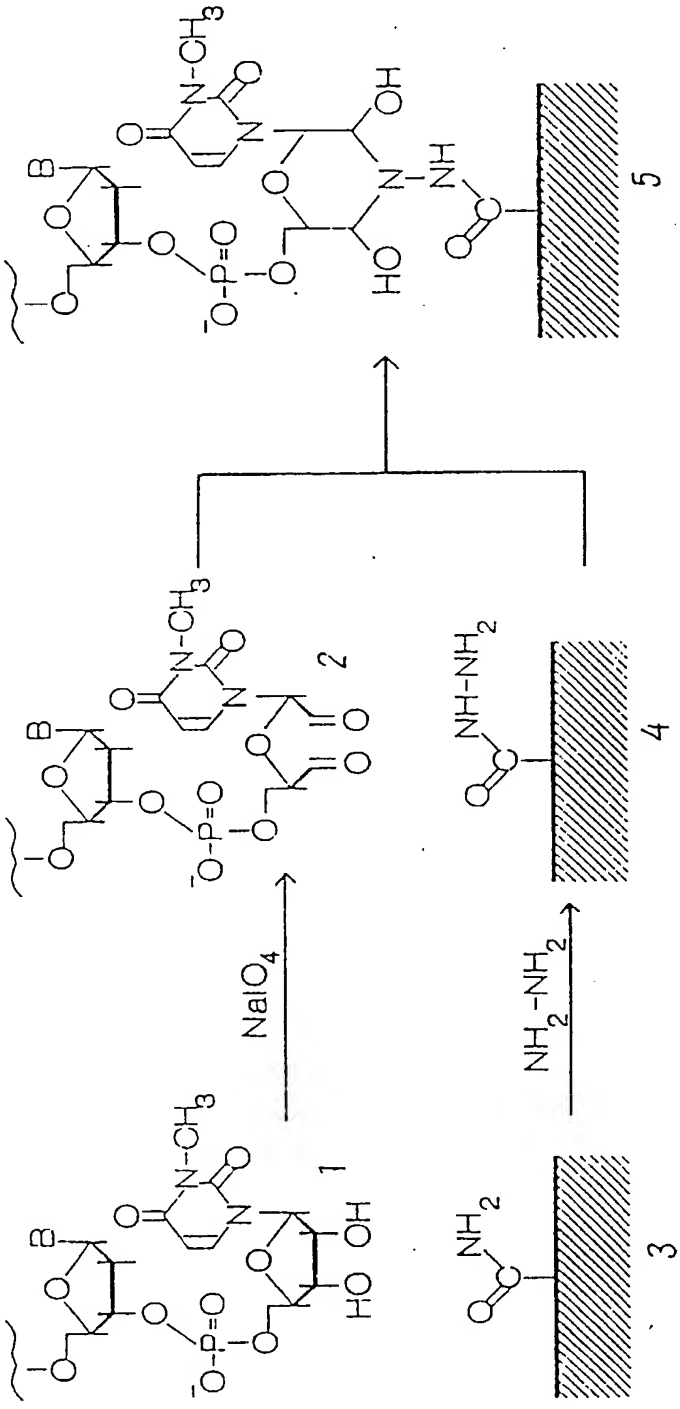


FIG.3

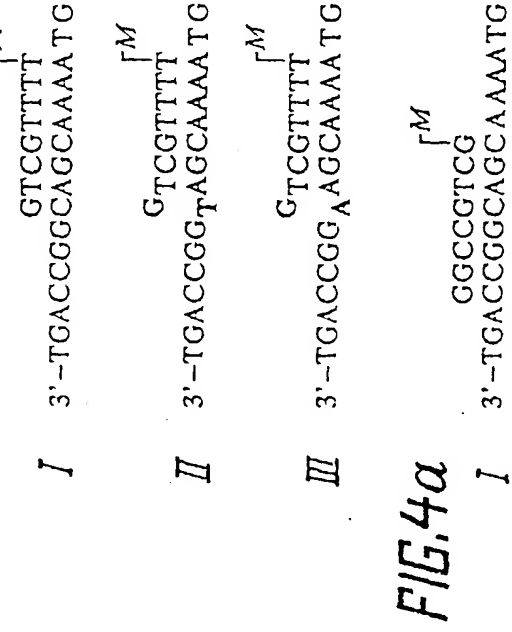
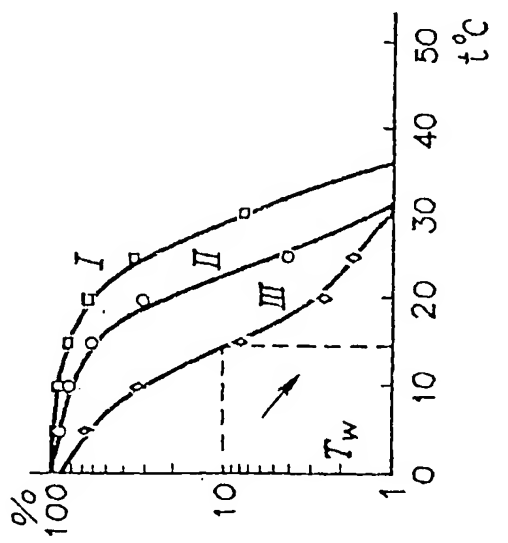

I 3'-TGACCGGCAGCAAAATG
II 3'-TGACCGGTAGCAAAATG
III 3'-TGACCGGAAGCAAAATG

FIG.4a


%
100
10
1
0
10
20
30
40
50
t°C
I
II
III
T_w

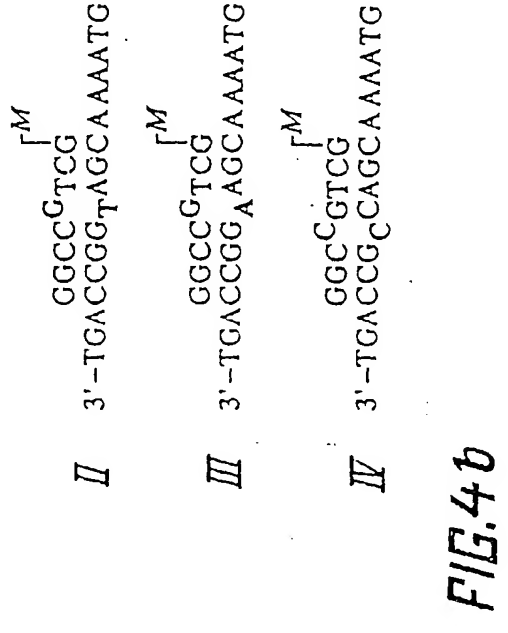
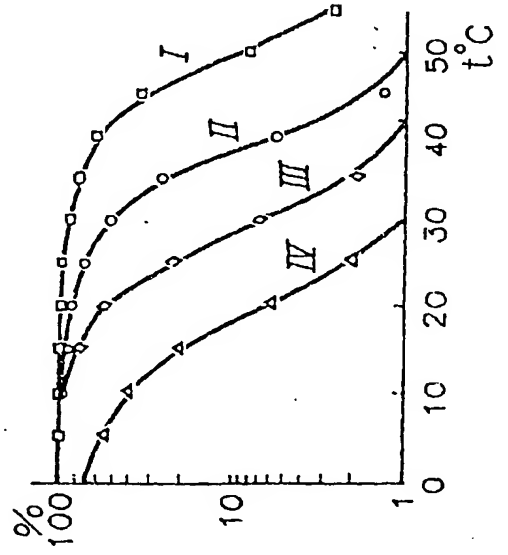

I 3'-TGACCGGCAGCAAAATG
II 3'-TGACCGGTAGCAAAATG
III 3'-TGACCGGAAGCAAAATG
IV 3'-TGACCGGCAGCAAAATG

FIG.4b


%
100
10
1
0
10
20
30
40
50
t°C
I
II
III
IV

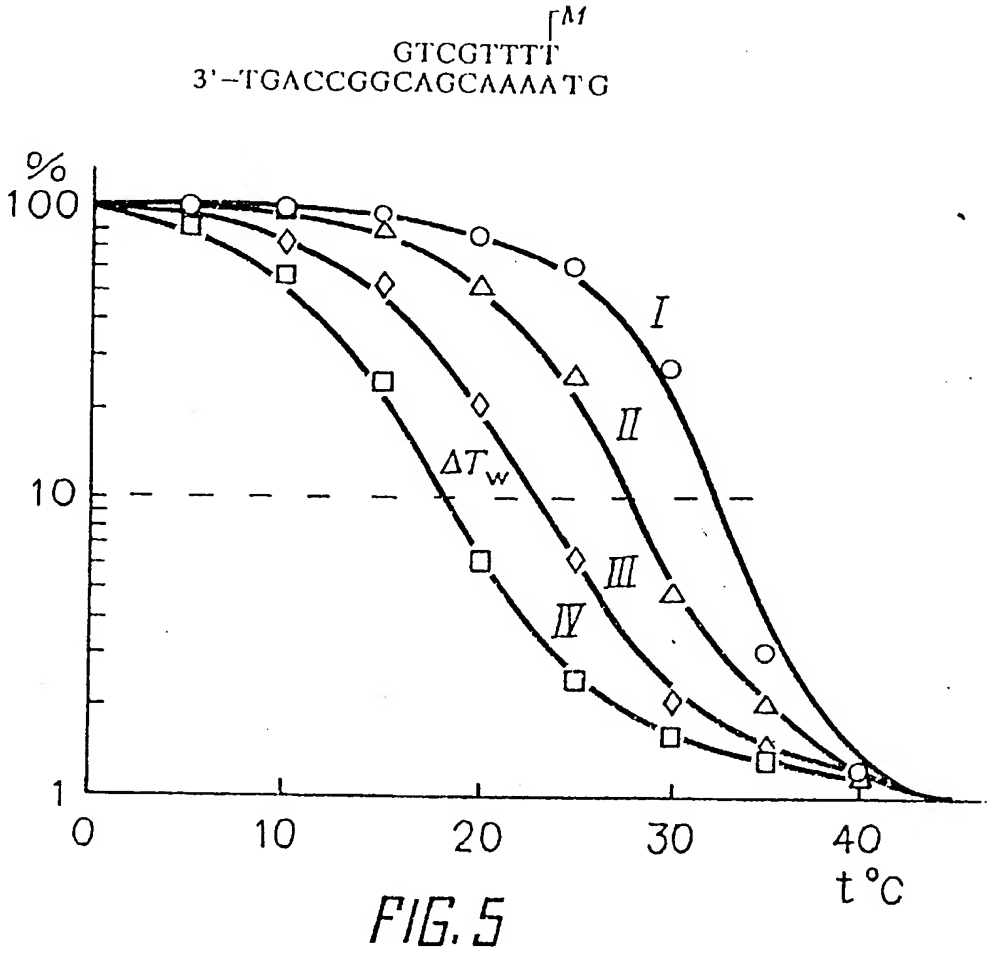

M1
GTCGTTTT
3'–TGACCGGCAGCAAAATG
%
100
10
1
0
10
20
30
40
t °C
I
II
III
IV
ΔTw

FIG. 5

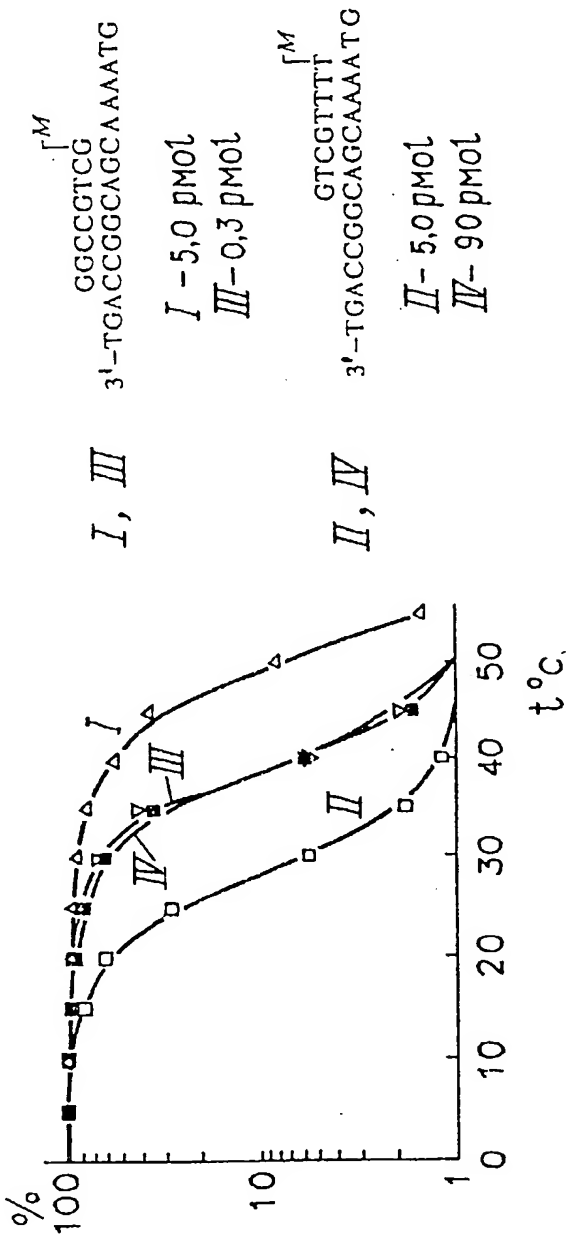

GGCCGTCG
3'-TGACCGGCAGCAAAATG
I, III
I - 5,0 pMol
III - 0,3 pMol
GTCGTTTT
3'-TGACCGGCAGCAAAATG
II, IV
II - 5,0 pMol
IV - 90 pMol
%
100
10
1
0
10
20
30
40
50
t °C
I
II
III
IV

FIG.6

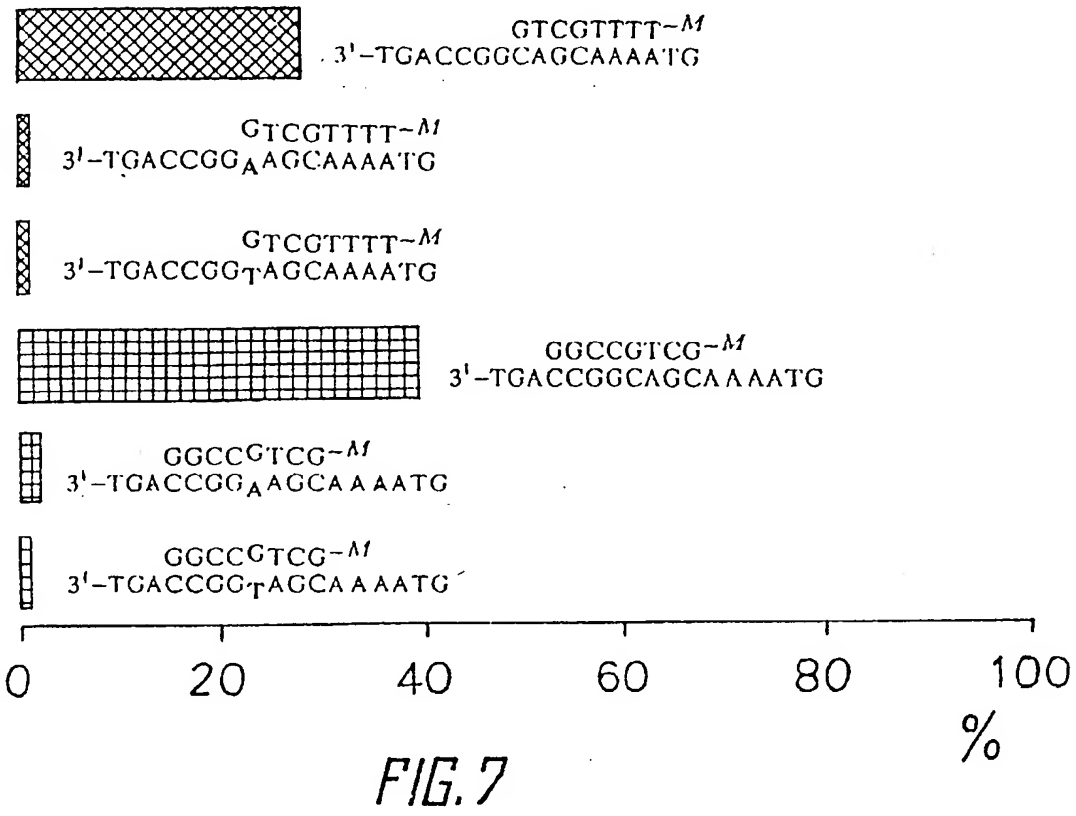

GTCGTTTT–M
3'–TGACCGGCAGCAAAATG
GTCGTTTT–M
3'–TGACCGGAAGCAAAATG
GTCGTTTT–M
3'–TGACCGGTAGCAAAATG
GGCCGTCG–M
3'–TGACCGGCAGCAAAATG
GGCCGTCG–M
3'–TGACCGGAAGCAAAATG
GGCCGTCG–M
3'–TGACCGGTAGCAAAATG
0
20
40
60
80
100
%

FIG. 7

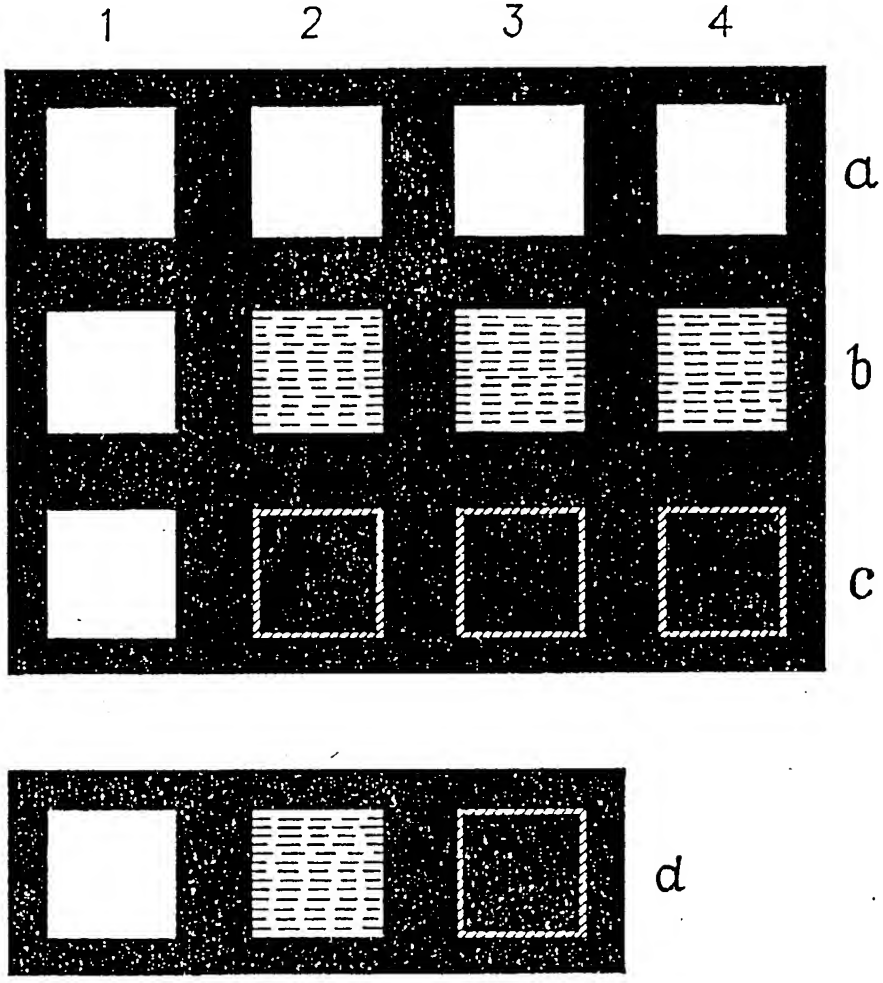


FIG. 8